DOI: 10.1214/ECP.v5-1025

Elect. Comm. in Probab. **5** (2000) 101–117

ELECTRONIC COMMUNICATIONS in PROBABILITY

A CONVERSE COMPARISON THEOREM FOR BSDES AND RELATED PROPERTIES OF g–EXPECTATION

PHILIPPE BRIAND
IRMAR, Université Rennes 1
35 042 Rennes Cedex, FRANCE
email: `Philippe.Briand@univ-rennes1.fr`

FRANÇOIS COQUET
IRMAR, Université Rennes 1
35 042 Rennes Cedex, FRANCE
email: `Francois.Coquet@univ-rennes1.fr`

YING HU
IRMAR, Université Rennes 1
35 042 Rennes Cedex, FRANCE
email: `Ying.Hu@univ-rennes1.fr`

JEAN MÉMIN
IRMAR, Université Rennes 1
35 042 Rennes Cedex, FRANCE
email: `Jean.Memin@univ-rennes1.fr`

SHIGE PENG
Department of Mathematics, Shandong University
Jinan 250100, CHINA
email: `peng@public.jn.sd.cn`

submitted October 9, 1999 *Final version accepted* May 23, 2000

AMS 2000 Subject classification: 60H10, 60H30
backward stochastic differential equations, comparison theorem

Abstract

In [1], Z. Chen *proved that, if for each terminal condition ξ, the solution of the BSDE associated to the standard parameter (ξ, g_1) is equal at time $t = 0$ to the solution of the BSDE associated to (ξ, g_2) then we must have $g_1 \equiv g_2$. This result yields a natural question: what happens in the case of an inequality in place of an equality? In this paper, we try to investigate this question and we prove some properties of "g–expectation", notion introduced by* S. Peng *in [8].*

1 Introduction

It is by now well-known that there exists a unique, adapted and square integrable, solution to a backward stochastic differential equation (BSDE for short in the remaining of the paper) of type

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s \cdot dW_s, \qquad 0 \le t \le T,$$

providing, for instance, that the generator is Lipschitz in both variables y and z and that ξ and $\big(f(s,0,0)\big)_{s\in[0,T]}$ are square integrable. We refer of course to E. Pardoux and S. Peng [4, 5] and to N. El Karoui, S. Peng and M.-C. Quenez [2] for a survey of the applications of this theory in finance.

One of the great achievement of the theory of BSDEs is the comparison theorem for real-valued BSDEs due to S. Peng [7] at first and then generalized by several authors, see e.g. N. El Karoui, S. Peng and M.-C. Quenez [2, Theorem 2.2]. It allows to compare the solutions of two BSDEs whenever we can compare the terminal conditions and the generators. In this paper we try to investigate an inverse problem: if we can compare the solutions of two BSDEs (at time $t = 0$) with the same terminal condition, for all terminal conditions, can we compare the generators?

The result of Z. Chen [1] can be read as the first step in solving this problem. Indeed, he proved using the language of "g–expectation" introduced by S. Peng in [8] that, given two generators, say g_1 and g_2, then if, for each $\xi \in \mathrm{L}^2$, we have $Y_0^1(\xi) = Y_0^2(\xi)$ where $\big((Y_t^i(\xi), Z_t^i(\xi)\big)_{t\in[0,T]}$ stands for the solution of the BSDE:

$$Y_t = \xi + \int_t^T g_i(s, Y_s, Z_s)ds - \int_t^T Z_s \cdot dW_s, \qquad 0 \le t \le T,$$

for $i = 1,\ 2$, then we have $g_1 \equiv g_2$. With the above notations, the main issue of this paper is to address the following question: if, for each $\xi \in \mathrm{L}^2$, we have $Y_0^1(\xi) \le Y_0^2(\xi)$, do we have $g_1 \le g_2$?

The paper is organized as follows: in section 2, we introduce some notations and we make our assumptions. In section 3, we prove the result in the case of deterministic generators and we give an application of these techniques to partial differential equations (PDEs for short in the rest of the paper). In section 4, we prove a converse to the comparison theorem for BSDEs and then, with the help of this result we study the case when the generators do not depend on the variable y. Finally, in section 5, we discuss the Jensen inequality for "g–expectation".

2 Preliminaries

2.1 Notations and assumptions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space carrying a standard d-dimensional Brownian motion, $(W_t)_{t\ge0}$, starting from $W_0 = 0$, and let $\big(\mathcal{F}_t\big)_{t\ge0}$ be the σ-algebra generated by $(W_t)_{t\ge0}$. We do the usual $\mathbb{P}$-augmentation to each $\mathcal{F}_t$ such that $\big(\mathcal{F}_t\big)_{t\ge0}$ is right continuous and complete. If z belongs to $\mathbb{R}^d$, $\|z\|$ denotes its Euclidean norm. We define the following usual spaces of processes:

- $\mathcal{S}_2 = \Big\{\psi \text{ progressively measurable;} \quad \|\psi\|^2_{\mathcal{S}_2} := \mathbb{E}\Big[\sup_{0\le t\le T} |\psi_t|^2\Big] < \infty\Big\}$,

- $\mathcal{H}_2 = \left\{ \psi \text{ progressively measurable}; \quad \|\psi\|_2^2 := \mathbb{E}\left[\int_0^T \|\psi_t\|^2 dt \right] < \infty \right\}.$

Let us consider a function g, which will be in the following the generator of the BSDE, defined on $\Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d$, with values in $\mathbb{R}$, s.t. the process $\big(g(t,y,z)\big)_{t\in[0,T]}$ is progressively measurable for each (y,z) in $\mathbb{R} \times \mathbb{R}^d$. For the function g, we will use, through-out the paper, the following assumptions.

(A 1). There exists a constant $K \geq 0$ s.t. $\mathbb{P} - a.s.$, we have:

$$\forall t,\ \forall (y,y'),\ \forall (z,z'),\quad \big|g(t,y,z) - g(t,y',z')\big| \leq K\big(|y-y'| + \|z-z'\|\big).$$

(A 2). The process $\big(g(t,0,0)\big)_{t\in[0,T]}$ belongs to $\mathcal{H}_2$.

(A 3). $\mathbb{P} - a.s., \forall (t,y), \quad g(t,y,0) = 0.$

(A 4). $\mathbb{P} - a.s., \forall (y,z), \quad t \longmapsto g(t,y,z)$ is continuous.

It is by now well known that under the assumptions (A 1) and (A 2), for any random variable ξ in $\mathrm{L}^2(\mathcal{F}_T)$, the BSDE

$$Y_t = \xi + \int_t^T g(s,Y_s,Z_s)ds - \int_t^T Z_s \cdot dW_s, \qquad 0 \leq t \leq T, \tag{1}$$

has a unique adapted solution, say $\big((Y_t,Z_t)\big)_{t\in[0,T]}$ s.t. Z is in the space $\mathcal{H}_2$. Actually, by classical results in this field, the process Y belongs to $\mathcal{S}_2$.
In [8], S. Peng adopted a new point of view in the study of BSDEs. Indeed, if g satisfies the assumptions (A 1) and (A 3), he introduced the function $\mathcal{E}_g$ defined on $\mathrm{L}^2(\mathcal{F}_T)$ with values in $\mathbb{R}$ by simply setting $\mathcal{E}_g(\xi) := Y_0$ where (Y,Z) is the solution of the BSDE (1) (since this solution is adapted, Y_0 is deterministic). He called $\mathcal{E}_g$ g–expectation and he proved that some properties of the classical expectation are preserved (monotonicity for instance) but, since g is not linear in general, the linearity is not preserved (we will see at the end of the paper that the Jensen inequality does not hold in general for $\mathcal{E}_g$).
Related to this "nonlinear expectation", he defined also a conditional g–expectation by setting $\mathcal{E}_g\big(\xi \,|\, \mathcal{F}_t\big) := Y_t$ which is the unique random variable η, $\mathcal{F}_t$–measurable and square–integrable, s.t.

$$\forall A \in \mathcal{F}_t, \qquad \mathcal{E}_g\big(\mathbf{1}_A \eta\big) = \mathcal{E}_g\big(\mathbf{1}_A \xi\big).$$

Z. Chen in [1] used these notions to prove the following result:

Theorem 2.1 *Let the assumptions (A 1), (A 3) and (A 4) hold for g_1 and g_2 and let us assume moreover that, for each $\xi \in \mathrm{L}^2(\mathcal{F}_T)$, $\mathcal{E}_{g_1}\big(\xi\big) = \mathcal{E}_{g_2}\big(\xi\big)$.*
Then, $\mathbb{P} - a.s., \forall (t,y,z), \quad g_1(t,y,z) = g_2(t,y,z)$.

In this note, as mentioned in the introduction, we will try to investigate the following question: if, for each $\xi \in \mathrm{L}^2(\mathcal{F}_T)$, $\mathcal{E}_{g_1}\big(\xi\big) \leq \mathcal{E}_{g_2}\big(\xi\big)$, do we have $g_1(t,y,z) \leq g_2(t,y,z)$? This problem is roughly speaking a converse to the comparison theorem for BSDEs, since if $g_1 \leq g_2$ then $\mathcal{E}_{g_1}(\cdot) \leq \mathcal{E}_{g_2}(\cdot)$.
We close this subsection by a comment about the assumption (A 3). Under (A 3), if ξ is a random variable $\mathcal{F}_S$–measurable with $S < T$, then we have: $\mathcal{E}_g\big(\xi \mid \mathcal{F}_t\big) = \xi$ if $S \leq t \leq T$ and

$\mathcal{E}_g\big(\xi \mid \mathcal{F}_t\big) = y_t$ for $0 \le t < S$ where $\big((y_r, z_r)\big)_{r\in[0,S]}$ stands for the solution on $[0, S]$ of the BSDE:

$$y_r = \xi + \int_r^S g(u, y_u, z_u)du - \int_r^S z_u \cdot dW_u, \qquad 0 \le r \le S.$$

This remark will be used in the proofs of Theorem 3.2, Theorem 4.1 and Lemma 4.3.

2.2 Technical Results

In this subsection, we establish a technical result which will be useful in the next section. We start by giving an a priori estimate for BSDEs which is of standard type, see [2].

Proposition 2.2 *Let $\xi \in \mathrm{L}^2(\mathcal{F}_T)$ and let the assumptions (A 1) and (A 2) hold. Then, if the process $\big((Y_t, Z_t)\big)_{t\in[0,T]}$ is the solution of the BSDE* (1) *we have, for $\beta = 2(K + K^2)$:*

$$\mathbb{E}\Big\{\sup_{t\le s\le T} \mathrm{e}^{\beta s}|Y_s|^2 + \int_t^T \mathrm{e}^{\beta s}\|Z_s\|^2 ds \;\Big|\; \mathcal{F}_t\Big\} \le C\mathbb{E}\Big\{\mathrm{e}^{\beta T}|\xi|^2 + \Big(\int_t^T \mathrm{e}^{(\beta/2)s}\big|g(s,0,0)\big|ds\Big)^2 \;\Big|\; \mathcal{F}_t\Big\},$$

where C is a universal constant.

Proof. We outline the proof for the convenience of the reader.

As usual we start with Itô's formula to see that, noting M_u for $2\int_u^T e^{\beta s} Y_s Z_s \cdot dW_s$, for each $u \in [0, T]$,

$$e^{\beta u}|Y_u|^2 + \int_u^T e^{\beta s}\|Z_s\|^2 ds = e^{\beta T}|\xi|^2 + 2\int_u^T e^{\beta s} Y_s \cdot g(s, Y_s, Z_s)ds - \int_u^T \beta e^{\beta s}|Y_s|^2 ds - M_u.$$

Using the Lipschitz assumption on g and then the inequality $2K|y| \cdot \|z\| \le 2K^2|y|^2 + (1/2)\|z\|^2$, we deduce that, taking into account the definition of β,

$$e^{\beta u}|Y_u|^2 + \frac{1}{2}\int_u^T e^{\beta s}\|Z_s\|^2 ds \le e^{\beta T}|\xi|^2 + 2\int_u^T e^{\beta s}|Y_s| \cdot |g(s,0,0)|ds - 2\int_u^T e^{\beta s} Y_s Z_s \cdot dW_s. \quad (2)$$

In particular, taking the conditional expectation w.r.t. $\mathcal{F}_t$ of the previous inequality written for $u = t$, we deduce since the conditional expectation of the stochastic integral vanishes,

$$\mathbb{E}\Big\{\int_t^T e^{\beta s}\|Z_s\|^2 ds \;\Big|\; \mathcal{F}_t\Big\} \le 2\mathbb{E}\Big\{e^{\beta T}|\xi|^2 + 2\int_t^T e^{\beta s}|Y_s| \cdot |g(s,0,0)|ds \;\Big|\; \mathcal{F}_t\Big\}. \quad (3)$$

Moreover, coming back to the inequality (2), we get

$$\sup_{t\le u\le T} e^{\beta u}|Y_u|^2 \le e^{\beta T}|\xi|^2 + 2\int_t^T e^{\beta s}|Y_s| \cdot |g(s,0,0)|ds + 4\sup_{t\le u\le T}\Big|\int_t^u e^{\beta s} Y_s Z_s \cdot dW_s\Big|,$$

and by Burkholder-Davis-Gundy's inequality, the previous estimate yields the inequality

$$\begin{aligned}\mathbb{E}\Big\{\sup_{t\le u\le T} e^{\beta u}|Y_u|^2 \;\Big|\; \mathcal{F}_t\Big\} \;&\le\; \mathbb{E}\Big\{e^{\beta T}|\xi|^2 + 2\int_t^T e^{\beta s}|Y_s| \cdot |g(s,0,0)|ds \;\Big|\; \mathcal{F}_t\Big\} \\ &\quad + C\mathbb{E}\Big\{\Big(\int_t^T e^{2\beta s}|Y_s|^2\|Z_s\|^2 ds\Big)^{1/2} \;\Big|\; \mathcal{F}_t\Big\},\end{aligned}$$

where C is a universal constant which will change from line to line; and thanks to the inequality $ab \le a^2/2 + b^2/2$ we deduce immediately

$$\begin{aligned}\mathbb{E}\Big\{ \sup_{t\le u\le T} e^{\beta u}|Y_u|^2 \,\Big|\, \mathcal{F}_t\Big\} \quad &\le \quad \mathbb{E}\Big\{ e^{\beta T}|\xi|^2 + 2\int_t^T e^{\beta s}|Y_s|\cdot|g(s,0,0)|ds \,\Big|\, \mathcal{F}_t\Big\} \\ &\quad + \frac{C^2}{2}\mathbb{E}\Big\{ \int_t^T e^{\beta s}\|Z_s\|^2 ds \,\Big|\, \mathcal{F}_t\Big\} + \frac{1}{2}\mathbb{E}\Big\{ \sup_{t\le u\le T} e^{\beta u}|Y_u|^2 \,\Big|\, \mathcal{F}_t\Big\}.\end{aligned}$$

Combining the inequality (3) with the previous one, we easily derive that, for a universal constant C, we have

$$\mathbb{E}\Big\{ \sup_{t\le u\le T} e^{\beta u}|Y_u|^2 + \int_t^T e^{\beta s}\|Z_s\|^2 ds \,\Big|\, \mathcal{F}_t\Big\} \le C\mathbb{E}\Big\{ e^{\beta T}|\xi|^2 + \int_t^T e^{\beta s}|Y_s|\cdot|g(s,0,0)|ds \,\Big|\, \mathcal{F}_t\Big\}.$$

It remains to use the fact that

$$\begin{aligned}C\mathbb{E}\Big\{ \int_t^T e^{\beta s}|Y_s|\cdot|g(s,0,0)|ds \,\Big|\, \mathcal{F}_t\Big\} \quad &\le \quad \frac{1}{2}\mathbb{E}\Big\{ \sup_{t\le u\le T} e^{\beta u}|Y_u|^2 \,\Big|\, \mathcal{F}_t\Big\} \\ &\quad + \frac{C^2}{2}\mathbb{E}\Big\{ \Big(\int_t^T e^{(\beta/2)s}|g(s,0,0)|ds\Big)^2 \,\Big|\, \mathcal{F}_t\Big\}\end{aligned}$$

and to change C one more time to finish the proof of this proposition. □

Before stating our first result, we need some further notations.

(A 5). Let $b : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \longrightarrow \mathbb{R}^{n\times d}$ be two Lipschitz functions.

We denote by $X^{t,x}$ the solution of the SDE:

$$X_s^{t,x} = x + \int_t^s b(X_u^{t,x})du + \int_t^s \sigma(X_u^{t,x})\cdot dW_u, \quad s \ge t, \tag{4}$$

with the usual convention $X_s^{t,x} = x$ if $s < t$.
Let $g : \Omega\times[0,T]\times\mathbb{R}\times\mathbb{R}^d \longrightarrow \mathbb{R}$ be a function s.t. the stochastic process $\big(g(s,y,z)\big)_{s\in[0,T]}$ is progressively measurable for each (y,z). Fix $(t,x,y,p) \in [0,T[\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^n$ and let us denote by $\big(({}^{\varepsilon}Y_s^{t,x,y,p}, {}^{\varepsilon}Z_s^{t,x,y,p})\big)_{s\in[0,t+\varepsilon]}$, for $\varepsilon > 0$ small enough, the solution of the BSDE on $[0,t+\varepsilon]$:

$${}^{\varepsilon}Y_s^{t,x,y,p} = y + p\cdot\big(X_{t+\varepsilon}^{t,x} - x\big) + \int_s^{t+\varepsilon} g(u, {}^{\varepsilon}Y_u^{t,x,y,p}, {}^{\varepsilon}Z_u^{t,x,y,p})du - \int_s^{t+\varepsilon} {}^{\varepsilon}Z_u^{t,x,y,p}\cdot dW_u. \tag{5}$$

We claim the following proposition:

Proposition 2.3 *Let the assumptions (A 1), (A 2) and (A 4) hold for the function g and let the notation (A 5) holds. Let us assume moreover that $\mathbb{E}\big[\sup_{0\le t\le T}|g(t,0,0)|^2\big]$ is finite. Then, for each $(t,x,y,p) \in [0,T[\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^n$, we have:*

$$\mathrm{L}^2 - \lim_{\varepsilon\to 0^+} \frac{1}{\varepsilon}\big\{ {}^{\varepsilon}Y_t^{t,x,y,p} - y\big\} = g(t,y,\sigma^t(x)p) + p\cdot b(x).$$

Proof. Since $(t,x,y,p)\in[0,T[\times\mathbb{R}^n\times\mathbb{R}\times\mathbb{R}^n$ is fixed in this proof, we will drop, for notational convenience, the superscript (t,x,y,p) and then write $(Y^\varepsilon,Z^\varepsilon)$ instead of $({}^\varepsilon Y^{t,x,y,p},{}^\varepsilon Z^{t,x,y,p})$. Firstly, let us remark that by classical results on SDEs, see e.g. H. Kunita [3], the terminal condition of the BSDE (5) is square integrable and then this BSDE has a unique solution in the space $\mathcal{S}_2\times\mathcal{H}_2$.
Let us pick $\varepsilon>0$ and let us define, for $t\le s\le t+\varepsilon$,

$$\widetilde{Y}^\varepsilon_s=Y^\varepsilon_s-\big(y+p\cdot(X^{t,x}_s-x)\big),\quad\text{and,}\quad \widetilde{Z}^\varepsilon_s=Z^\varepsilon_s-\sigma^t(X^{t,x}_s)p.$$

Using Itô's formula, we easily show that, on the time interval $[t,t+\varepsilon]$, the process $(\widetilde{Y}^\varepsilon,\widetilde{Z}^\varepsilon)$ solves the following BSDE:

$$\widetilde{Y}^\varepsilon_s=\int_s^{t+\varepsilon}g\big(u,y+p\cdot(X^{t,x}_u-x)+\widetilde{Y}^\varepsilon_u,\sigma^t(X^{t,x}_u)p+\widetilde{Z}^\varepsilon_u\big)du+\int_s^{t+\varepsilon}p\cdot b(X^{t,x}_u)du-\int_s^{t+\varepsilon}\widetilde{Z}^\varepsilon_u\cdot dW_u. \quad (6)$$

Thus, by Proposition 2.2, we have the following estimate:

$$\begin{aligned}\mathbb{E}\Big\{\sup_{t\le s\le t+\varepsilon}|\widetilde{Y}^\varepsilon_s|^2+\int_t^{t+\varepsilon}\|\widetilde{Z}^\varepsilon_s\|^2ds\,\Big|\,\mathcal{F}_t\Big\}\\ \le C\mathrm{e}^{2(K+K^2)\varepsilon}\mathbb{E}\Big\{\Big(\int_t^{t+\varepsilon}\big|g\big(u,y+p\cdot(X^{t,x}_u-x),\sigma^t(X^{t,x}_u)p\big)+p\cdot b(X^{t,x}_u)\big|du\Big)^2\,\Big|\,\mathcal{F}_t\Big\},\end{aligned}$$

and then, since g, b and σ are Lipschitz, we deduce that

$$\mathbb{E}\Big\{\sup_{t\le s\le t+\varepsilon}|\widetilde{Y}^\varepsilon_s|^2+\int_t^{t+\varepsilon}\|\widetilde{Z}^\varepsilon_s\|^2ds\,\Big|\,\mathcal{F}_t\Big\}\le C_{x,y,p}\varepsilon^2\mathbb{E}\Big\{1+\sup_{t\le s\le t+\varepsilon}\Big(|X^{t,x}_s|^2+|g(s,0,0)|^2\Big)\,\Big|\,\mathcal{F}_t\Big\}.$$

Taking the expectation in the previous inequality we obtain, since, by the additional assumption, $\mathbb{E}\big[\sup_{0\le s\le T}\big(|X^{t,x}_s|^2+|g(s,0,0)|^2\big)\big]\le C(1+|x|^2)$, the following estimate

$$\mathbb{E}\Big\{\sup_{t\le s\le t+\varepsilon}|\widetilde{Y}^\varepsilon_s|^2+\int_t^{t+\varepsilon}\|\widetilde{Z}^\varepsilon_s\|^2ds\Big\}\le C\varepsilon^2, \quad (7)$$

where C depends on x,y,p which is not important here since (x,y,p) are fixed.
With this inequality in hands, it is easy to prove the result. Indeed, taking the conditional expectation in the BSDE (6), we get

$$\frac{1}{\varepsilon}\big\{Y^\varepsilon_t-y\big\}=\frac{1}{\varepsilon}\widetilde{Y}^\varepsilon_t=\frac{1}{\varepsilon}\mathbb{E}\Big\{\int_t^{t+\varepsilon}\big\{g\big(u,y+p\cdot(X^{t,x}_u-x)+\widetilde{Y}^\varepsilon_u,\sigma^t(X^{t,x}_u)p+\widetilde{Z}^\varepsilon_u\big)+p\cdot b(X^{t,x}_u)\big\}du\Big|\,\mathcal{F}_t\Big\}.$$

We split the right hand side of the previous equality as follows:

$$\frac{1}{\varepsilon}\big\{Y^\varepsilon_t-y\big\}=\frac{1}{\varepsilon}\mathbb{E}\Big\{\int_t^{t+\varepsilon}\big\{g\big(u,y+p\cdot(X^{t,x}_u-x),\sigma^t(X^{t,x}_u)p\big)+p\cdot b(X^{t,x}_u)\big\}du\,\Big|\,\mathcal{F}_t\Big\}+R_\varepsilon,$$

where R_ε stands for

$$\frac{1}{\varepsilon}\mathbb{E}\Big\{\int_t^{t+\varepsilon}\big\{g\big(u,y+p\cdot(X^{t,x}_u-x)+\widetilde{Y}^\varepsilon_u,\sigma^t(X^{t,x}_u)p+\widetilde{Z}^\varepsilon_u\big)-g\big(u,y+p\cdot(X^{t,x}_u-x),\sigma^t(X^{t,x}_u)p\big)\big\}du\Big|\,\mathcal{F}_t\Big\}.$$

It is very easy to check that R_ε goes to 0 in L^2 when ε tends to 0^+. Indeed, since g is Lipschitz, we have

$$|R_\varepsilon| \le \frac{K}{\varepsilon}\mathbb{E}\bigg\{ \int_t^{t+\varepsilon} \{|\widetilde{Y}_u^\varepsilon| + \|\widetilde{Z}_u^\varepsilon\|\} du \,\bigg|\, \mathcal{F}_t \bigg\},$$

and then using Hölder's inequality we obtain

$$\mathbb{E}\Big[|R_\varepsilon|^2\Big] \le \frac{K^2}{\varepsilon^2}\mathbb{E}\bigg[\Big(\int_t^{t+\varepsilon} \{|\widetilde{Y}_u^\varepsilon| + \|\widetilde{Z}_u^\varepsilon\|\} du\Big)^2\bigg] \le \frac{2K^2}{\varepsilon}\mathbb{E}\bigg[\int_t^{t+\varepsilon} \{|\widetilde{Y}_u^\varepsilon|^2 + \|\widetilde{Z}_u^\varepsilon\|^2\} du\bigg].$$

Taking into account the estimate (7), the previous inequality yields $\mathbb{E}\big[|R_\varepsilon|^2\big] \le C(\varepsilon^2+\varepsilon)$ which shows the convergence of R_ε to 0.
It remains only to check that, as $\varepsilon \to 0^+$, in the sense of L^2,

$$\frac{1}{\varepsilon}\mathbb{E}\bigg\{ \int_t^{t+\varepsilon} \{g\big(u, y+p\cdot(X_u^{t,x}-x), \sigma^t(X_u^{t,x})p\big) + p\cdot b(X_u^{t,x})\} du \,\bigg|\, \mathcal{F}_t \bigg\} \longrightarrow g\big(t,y,\sigma^t(x)p\big) + p\cdot b(x).$$

For this task, we set

$$Q_\varepsilon = \frac{1}{\varepsilon}\mathbb{E}\bigg\{ \int_t^{t+\varepsilon} \Big\{g\big(u, y+p\cdot(X_u^{t,x}-x), \sigma^t(X_u^{t,x})p\big) - g\big(u,y,\sigma^t(x)p\big) + p\cdot\big(b(X_u^{t,x})-b(x)\big)\Big\} du \,\bigg|\, \mathcal{F}_t \bigg\},$$

and we will prove that $Q_\varepsilon \longrightarrow 0$ in L^2 and that the same is true for the remaining term, i.e.

$$P_\varepsilon = \frac{1}{\varepsilon}\mathbb{E}\bigg\{ \int_t^{t+\varepsilon} \Big\{g\big(u,y,\sigma^t(x)p\big) - g\big(t,y,\sigma^t(x)p\big)\Big\} du \,\bigg|\, \mathcal{F}_t \bigg\}.$$

For the first part, we remark that, since g, b and σ are Lipschitz,

$$|Q_\varepsilon| \le \frac{C}{\varepsilon}\mathbb{E}\bigg\{ \int_t^{t+\varepsilon} \big|X_u^{t,x} - x\big| du \,\bigg|\, \mathcal{F}_t \bigg\},$$

and then, by Hölder's inequality, we get

$$\mathbb{E}\Big[|Q_\varepsilon|^2\Big] \le \frac{C^2}{\varepsilon^2}\mathbb{E}\bigg[\Big(\int_t^{t+\varepsilon} \big|X_u^{t,x} - x\big| du\Big)^2\bigg] \le \frac{C^2}{\varepsilon}\mathbb{E}\bigg[\int_t^{t+\varepsilon} \big|X_u^{t,x} - x\big|^2 du\bigg].$$

But since we have (see e.g. H. Kunita [3]), for each $r \ge 1$,

$$\mathbb{E}\Big[\sup_{0\le s\le T} \big|X_s^{t,x}\big|^r\Big] \le C_T\big(1+|x|^r\big),$$

the function $u \longmapsto \mathbb{E}\Big[\big|X_u^{t,x} - x\big|^2\Big]$ is continuous. This function is equal to 0 at time t, from which we deduce that $Q_\varepsilon \longrightarrow 0$ in L^2.
Finally, we have, by Hölder's inequality,

$$\mathbb{E}\Big[|P_\varepsilon|^2\Big] \le \mathbb{E}\bigg[\frac{1}{\varepsilon}\int_t^{t+\varepsilon} \big|g\big(u,y,\sigma^t(x)p\big) - g\big(t,y,\sigma^t(x)p\big)\big|^2 du\bigg].$$

But by the assumption (A 4), $\displaystyle\frac{1}{\varepsilon}\int_t^{t+\varepsilon} \big|g\big(u,y,\sigma^t(x)p\big) - g\big(t,y,\sigma^t(x)p\big)\big|^2 du$ tends to 0 $\mathbb{P}$-a.s., and moreover

$$\frac{1}{\varepsilon}\int_t^{t+\varepsilon} \big|g\big(u,y,\sigma^t(x)p\big) - g\big(t,y,\sigma^t(x)p\big)\big|^2 du \le C\big(\sup_{0\le s\le T}|g(s,0,0)|^2 + |y|^2 + \|\sigma^t(x)p\|^2\big)$$

which is assumed to be integrable. Thus the result follows from Lebesgue's theorem.
The proof is complete. $\square$

Remark. As we can see in the proof, the continuity of the process $\big(g(t,y,z)\big)_{t\in[0,T]}$ (assumption (A 4)) is not really needed. We can prove the result if this process is only right-continuous.

Remark. It is worth noting that the assumption "$\mathbb{E}\big[\sup_{0\le t\le T}|g(t,0,0)|^2\big]$ finite" holds when g satisfies (A 3) or in the Markovian situation (see the subsection 3.2 below).

3 The deterministic case

3.1 Main result

This subsection is devoted to the study of the deterministic case for which we can prove a useful property of g–expectation. In this paragraph, g is defined from $[0,T]\times\mathbb{R}\times\mathbb{R}^d$ into $\mathbb{R}$ and satisfies (A 1) and (A 3). Since $\mathcal{E}_g$ is a generalization of the expectation, a natural question which arises is the following: if ξ is independent of $\mathcal{F}_t$, do we have $\mathcal{E}_g\big(\xi\,\big|\,\mathcal{F}_t\big)=\mathcal{E}_g\big(\xi\big)$? We claim the following proposition which is mainly contained in [2].

Proposition 3.1 *If g is deterministic, then, for each $\xi\in\mathrm{L}^2(\mathcal{F}_T)$, we have $\mathcal{E}_g\big(\xi\,\big|\,\mathcal{F}_t\big)=\mathcal{E}_g\big(\xi\big)$ as soon as ξ is independent of $\mathcal{F}_t$.*

Proof. It is enough to check, from the assumption (A 3), that $\mathcal{E}_g\big(\xi\,\big|\,\mathcal{F}_t\big)$ is deterministic. Indeed, by construction we have $\mathcal{E}_g\big(\xi\big)=\mathcal{E}_g\big\{\mathcal{E}_g\big(\xi\,\big|\,\mathcal{F}_t\big)\big\}$ and, under the assumption (A 3) (see S. Peng [8, Lemma 36.3]) for each constant c, $\mathcal{E}_g(c)=c$.
We use the shift method to see that $\mathcal{E}_g\big(\xi\,\big|\,\mathcal{F}_t\big)$ is deterministic. Let us introduce, for each s s.t. $0\le s\le T-t$, $W'_s=W_{t+s}-W_t$. Then $\big\{W'_s,\,0\le s\le T-t\big\}$ is a Brownian motion w.r.t. its filtration $\mathcal{F}'_s$ which is nothing but $\mathcal{F}^t_{t+s}$ the σ–algebra generated by the increments of the Brownian motion W after the time t. Since ξ is $\mathcal{F}_T$–measurable and independent of $\mathcal{F}_t$, ξ is measurable w.r.t. $\mathcal{F}'_{T-t}$. As a consequence, we can construct the solution $\big((Y'_s,Z'_s)\big)_{0\le s\le T-t}$ of the BSDE

$$Y'_s=\xi+\int_s^{T-t}g(t+u,Y'_u,Z'_u)du-\int_s^{T-t}Z'_u\cdot dW'_u,\qquad 0\le s\le T-t,$$

and, setting $s=v-t$, we get

$$Y'_{v-t}=\xi+\int_{v-t}^{T-t}g(t+u,Y'_u,Z'_u)du-\int_{v-t}^{T-t}Z'_u\cdot dW'_u,\qquad t\le v\le T.$$

Now, if we make the change of variables $r=t+u$ in the integrals, we deduce that

$$\begin{aligned}Y'_{v-t}&=\xi+\int_v^T g(r,Y'_{r-t},Z'_{r-t})dr-\int_v^T Z'_{r-t}\cdot dW'_{r-t},\qquad t\le v\le T,\\&=\xi+\int_v^T g(r,Y'_{r-t},Z'_{r-t})dr-\int_v^T Z'_{r-t}\cdot dW_r,\qquad t\le v\le T.\end{aligned}$$

It follows that $\big((Y'_{v-t},Z'_{v-t})\big)_{v\in[t,T]}$ is a solution of the BSDE (1) on the time interval $[t,T]$ and then by uniqueness $\big((Y'_{v-t},Z'_{v-t})\big)_{v\in[t,T]}=\big((Y_v,Z_v)\big)_{v\in[t,T]}$. In particular, $Y'_0=Y_t$, and since Y'_0 is deterministic, the proof is complete. $\square$

Remark. Actually, if for each random variable $\xi \in \mathrm{L}^2(\mathcal{F}_T)$, we have $\mathcal{E}_g\big(\xi \,\big|\, \mathcal{F}_t\big) = \mathcal{E}_g\big(\xi\big)$ as soon as ξ is independent of $\mathcal{F}_t$ then g is deterministic.
Indeed, for any (t,y,z), we have, by Proposition 2.3,

$$g(t,y,z) = \lim_{\varepsilon\to 0^+} \frac{1}{\varepsilon}\big\{\mathcal{E}_g\big(y + z\cdot(W_{t+\varepsilon} - W_t) \,\big|\, \mathcal{F}_t\big) - y\big\} = \lim_{\varepsilon\to 0^+} \frac{1}{\varepsilon}\big\{\mathcal{E}_g\big(y + z\cdot(W_{t+\varepsilon} - W_t)\big) - y\big\},$$

and thus g is deterministic.
We give a counterexample when g is not deterministic. Actually, even in the simplest case, the linear case which corresponds to a Girsanov change of measures, the above property does not hold.
To show that, let us fix $T > 0$ and let us pick $t \in]0,T[$. Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be a continuous and bounded function. We define, for each $(s,z) \in [0,T]\times\mathbb{R}$, $g(s,z) := f(W_{s\wedge t})z$. Moreover, let us set $\xi = W_T - W_t$ which of course is independent of $\mathcal{F}_t$. By classical results for linear BSDEs, see e.g. [2, Proposition 2.2],

$$\begin{aligned}\mathcal{E}_g\big(\xi \,\big|\, \mathcal{F}_t\big) &= \mathbb{E}\bigg(\xi \exp\Big\{\int_t^T f(W_{s\wedge t})dW_s - \frac{1}{2}\int_t^T f^2(W_{s\wedge t})ds\Big\} \,\Big|\, \mathcal{F}_t\bigg)\\ &= \mathbb{E}^{\mathbb{Q}}\Big(\big(W_T - W_t\big) \,\big|\, \mathcal{F}_t\Big),\end{aligned}$$

where $\mathbb{Q}$ is the probability measure on $\big(\Omega, \mathcal{F}_T\big)$ whose density w.r.t. $\mathbb{P}$ is given by

$$Z_T := \exp\Big\{\int_0^T f(W_{s\wedge t})dW_s - \frac{1}{2}\int_0^T f^2(W_{s\wedge t})ds\Big\}.$$

But under the measure $\mathbb{Q}$, $\Big\{\widetilde{W}_r = W_r - \displaystyle\int_0^r f(W_{s\wedge t})ds,\ 0 \le r \le T\Big\}$ is a Brownian motion.
Moreover $W_T - W_t = \widetilde{W}_T - \widetilde{W}_t + \int_t^T f(W_{s\wedge t})ds = \widetilde{W}_T - \widetilde{W}_t + (T-t)f(W_t)$.
It follows immediately that $\mathcal{E}_g\big(\xi \,\big|\, \mathcal{F}_t\big) = (T-t)f(W_t)$ and thus, if f is not constant, we see that $\mathcal{E}_g\big(W_T - W_t \,\big|\, \mathcal{F}_t\big)$ is not deterministic which gives the desired result.

We are now able to answer the question asked in the introduction in the deterministic context. Let, for $i = 1,2$, $g_i : [0,T]\times\mathbb{R}\times\mathbb{R}^d \longrightarrow \mathbb{R}$. We claim the following result:

Theorem 3.2 *Let the assumptions (A 1), (A 3) and (A 4) hold for g_i, $i = 1,2$. Assume moreover that,*

$$\forall \xi \in \mathrm{L}^2(\mathcal{F}_T), \qquad \mathcal{E}_{g_1}(\xi) \le \mathcal{E}_{g_2}(\xi).$$

Then, we have,

$$\forall t \in [0,T],\ \forall (y,z) \in \mathbb{R}\times\mathbb{R}^d, \qquad g_1(t,y,z) \le g_2(t,y,z).$$

Proof. Let us fix $(t,y,z) \in [0,T[\times\mathbb{R}\times\mathbb{R}^d$, and, for $n \in \mathbb{N}^*$ large enough, let us consider $\xi_n = y + z\cdot(W_{t+(1/n)} - W_t)$. Since the assumption (A 3) holds (see the remark after the proof of Proposition 2.3), we have, by Proposition 2.3,

$$n\big({}^iY_t^n - y\big) \longrightarrow g_i(t,y,z), \quad \text{as } n\to\infty, \qquad \text{in } \mathrm{L}^2,$$

where $\big({}^iY^n, {}^iZ^n\big)$ is the solution of the BSDE

$${}^iY_t^n = \xi_n + \int_t^T g_i(r, {}^iY_r^n, {}^iZ_r^n)dr - \int_t^T {}^iZ_r^n\cdot dW_r, \qquad 0 \le t \le T.$$

In other words,

$$n\{\mathcal{E}_{g_i}(\xi_n \mid \mathcal{F}_t) - y\} \longrightarrow g_i(t,y,z), \qquad \text{in } \mathrm{L}^2$$

On the other hand, since ξ_n is independent of $\mathcal{F}_t$, Proposition 3.1 ensures that, for $i=1,2$, $\mathcal{E}_{g_i}(\xi_n \mid \mathcal{F}_t) = \mathcal{E}_{g_i}(\xi_n)$ and then by the hypothesis we deduce that

$$n\{\mathcal{E}_{g_1}(\xi_n \mid \mathcal{F}_t) - y\} \le n\{\mathcal{E}_{g_2}(\xi_n \mid \mathcal{F}_t) - y\}.$$

Passing to the limit when n goes to infinity, we obtain, since g_1 and g_2 are deterministic, the inequality $g_1(t,y,z) \le g_2(t,y,z)$, which concludes the proof, since (t,y,z) is arbitrary and both functions $g_1(\cdot,y,z)$ and $g_2(\cdot,y,z)$ are continuous at the point T. □

Remark. As we can see in the proof, we only need to assume that $\mathcal{E}_{g_1}(\xi) \le \mathcal{E}_{g_2}(\xi)$ for ξ of the form $y + z\cdot(W_s - W_t)$, for each (t,s,y,z), to get the result of this theorem. Actually, we can weaken a little bit more the assumption since it is enough to have the property when $s-t$ is small enough, say less than δ, and this δ may depend on (y,z).

3.2 An application to PDEs

We give in this subsection an application of the techniques described before to partial differential equations (PDEs for short). Semilinear PDEs constitute one of the field of applications of the theory of BSDEs as it was revealed by S. Peng [6] in the classical case and by E. Pardoux and S. Peng [5] for viscosity solutions of PDEs. Our setting is very close to that of F. Pradeilles [10].
Let $f : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^d \longrightarrow \mathbb{R}$ be a function s.t.

(A 6). There exist two constants $K \ge 0$ and $q \ge 1$, s.t.

1. $\forall (y,z), \quad x \longrightarrow f(x,y,z)$ is continuous;
2. $\forall x,\ \forall (y,z),(y',z'), \quad \big|f(x,y,z) - f(x,y',z')\big| \le K\big(|y-y'| + \|z-z'\|\big)$;
3. $\forall x, \quad \big|f(x,0,0)\big| \le K\big(1+|x|^q\big)$.

In addition, let us consider $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ a continuous function with polynomial growth. Then it is by now well known that BSDEs in Markovian context give a nonlinear Feynman-Kac formula for semilinear PDEs in the sense that the viscosity solution, say u, of the semilinear PDE

$$\partial_t u = \frac{1}{2}\mathrm{tr}\big(\sigma\sigma^t(x)\mathrm{D}^2 u\big) + b(x)\cdot\nabla u + f(x,u,\sigma^t(x)\nabla u), \quad (t,x) \in \mathbb{R}_+ \times \mathbb{R}^n, \qquad u(0,\cdot) = h(\cdot), \tag{8}$$

is equal, for each (t,x), to $Y_0^{t,x}$ where $\big((Y_s^{t,x}, Z_s^{t,x})\big)_{s\in[0,t]}$ is the solution of the BSDE

$$Y_s^{t,x} = h\big(X_t^{0,x}\big) + \int_s^t f(X_u^{0,x}, Y_u^{t,x}, Z_u^{t,x})du - \int_s^t Z_u^{t,x}\cdot dW_u, \qquad 0 \le s \le t,$$

where $X^{0,x}$ is the solution of the SDE (4).
Using this formula, we claim the following proposition:

Proposition 3.3 *Assume that the assumption (A 6) holds for two functions f_1 and f_2 and that b and σ satisfies the assumption (A 5).*
If, for $(x,y,p) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n$, there exists $\eta > 0$ s.t. $\forall \varepsilon < \eta$, $u_1(\varepsilon,x) \le u_2(\varepsilon,x)$, where u_i is the viscosity solution of the PDE (8) with the semilinear part f_i and the initial condition equal to $h(\cdot) = y + p\cdot(\cdot - x)$, then $f_1(x,y,\sigma^t(x)p) \le f_2(x,y,\sigma^t(x)p)$.

Proof. Let us rewrite the assumption in a probabilistic way using the nonlinear Feynman-Kac formula mentioned above. Fix (x,y,p) in the space $\mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n$. Then, there exists $\eta > 0$, s.t. $\forall \varepsilon < \eta$, ${}^1Y_0^{\varepsilon,x} \le {}^2Y_0^{\varepsilon,x}$, where $({}^iY^{\varepsilon,x}, {}^iZ^{\varepsilon,x})$ is the solution of the BSDE

$$ {}^iY_s^{\varepsilon,x} = y + p\cdot\left(X_\varepsilon^{0,x} - x\right) + \int_s^\varepsilon f_i(X_u^{0,x}, {}^iY_u^{\varepsilon,x}, {}^iZ_u^{\varepsilon,x})du - \int_s^\varepsilon {}^iZ_u^{\varepsilon,x}\cdot dW_u, \qquad 0 \le s \le \varepsilon. $$

This solution exists since by classical results on SDEs, see e.g. H. Kunita [3], we have,

$$ \mathbb{E}\Big[\sup_{0\le s\le \eta}\big|X_s^{0,x}\big|^{2q}\Big] \le C_\eta\big(1+|x|^{2q}\big), $$

and then, by the growth assumption on f_i, $\mathbb{E}\Big[\sup_{0\le s\le\varepsilon}\big|f(X_s^{0,x},0,0)\big|^2\Big]$ is finite.
But, the assumption yields

$$ \forall\varepsilon < \eta, \qquad \frac{1}{\varepsilon}\big\{{}^1Y_0^{\varepsilon,x} - y\big\} \le \frac{1}{\varepsilon}\big\{{}^2Y_0^{\varepsilon,x} - y\big\}, $$

and, on the other hand, by Proposition 2.3, in the sense of L^2,

$$ \frac{1}{\varepsilon}\big\{{}^iY_0^{\varepsilon,x} - y\big\} \longrightarrow f_i(x,y,\sigma^t(x)p) + p\cdot b(x), \quad \text{as } \varepsilon \to 0^+. $$

Since moreover $\frac{1}{\varepsilon}\big\{{}^iY_0^{\varepsilon,x} - y\big\}$ is deterministic the above convergence holds in $\mathbb{R}$ which gives the result. □

Remark. We do not suppose in this result that $f_i(x,y,0) = 0$.

4 Converse comparison theorem and application

4.1 A converse to the comparison theorem for BSDEs

Since the work of S. Peng [7], it is known that, under the usual assumptions, if $\mathbb{P} - a.s.$, the inequality $g_1(t,y,z) \le g_2(s,y,z)$ holds, then for each ξ, we have

$$ \forall t \in [0,T], \qquad Y_t^1(\xi) \le Y_t^2(\xi), $$

where $\big((Y_s^i(\xi), Z_s^i(\xi))\big)_{0\le s\le T}$ is the solution of the BSDE:

$$ Y_t^i(\xi) = \xi + \int_t^T g_i\big(s, Y_s^i(\xi), Z_s^i(\xi)\big)ds - \int_t^T Z_s^i(\xi)\cdot dW_s, \qquad 0 \le t \le T. $$

We prove in the next theorem a converse to this result with the additional assumption (A 3).

Theorem 4.1 *Let the assumptions (A 1), (A 3) and (A 4) hold for g_i, $i = 1,2$. Assume moreover that,*

$$\forall \xi \in \mathrm{L}^2(\mathcal{F}_T), \quad \forall t \in [0,T], \qquad Y_t^1(\xi) \le Y_t^2(\xi),$$

then, we have, $\mathbb{P}-a.s.$,

$$\forall t \in [0,T],\ \forall (y,z) \in \mathbb{R}\times\mathbb{R}^d, \qquad g_1(t,y,z) \le g_2(t,y,z).$$

Proof. Let us fix $(t,y,z) \in [0,T[\times\mathbb{R}\times\mathbb{R}^d$, and, for $n \in \mathbb{N}^*$ large enough, let us consider $\xi_n = y + z\cdot(W_{t+(1/n)} - W_t)$. We have by Proposition 2.3

$$n\big(Y_t^i(\xi_n) - y\big) \longrightarrow g_i(t,y,z), \qquad \text{in } \mathrm{L}^2.$$

On the other hand, by the hypothesis we deduce that

$$n\big\{Y_t^1(\xi_n) - y\big\} \le n\big\{Y_t^2(\xi_n) - y\big\}.$$

Extracting a subsequence to get the convergence $\mathbb{P}-a.s.$ and then passing to the limit when n goes to infinity, we obtain, the inequality $\mathbb{P}-a.s.$, $g_1(t,y,z) \le g_2(t,y,z)$. By continuity, we obtain finally that $\mathbb{P}-a.s.$,

$$\forall (t,y,z) \in [0,T]\times\mathbb{R}\times\mathbb{R}^d, \qquad g_1(t,y,z) \le g_2(t,y,z).$$

The proof is complete. □

Remark. We give in this remark the main lines of a different proof of this result. This approach is based on a "nonlinear decomposition theorem of Doob–Meyer's type" due to S. Peng [9]. For a given ξ, the assumption of the theorem says that the process $Y^2(\xi)$ is a g_1–supermartingale and actually it can be seen that it is a g_1–supermartingale in the strong sense. Therefore, we can apply Theorem 3.3 in [9] to see that this process is also a g_1–supersolution. From this we deduce easily that, for each ξ and for each $0 \le t_0 < t \le T$,

$$\frac{1}{t-t_0}\int_{t_0}^t \mathbb{E}\Big\{g_1\big(u, Y_u^2(\xi), Z_u^2(\xi)\big)\Big\}du \le \frac{1}{t-t_0}\int_{t_0}^t \mathbb{E}\Big\{g_2\big(u, Y_u^2(\xi), Z_u^2(\xi)\big)\Big\}du,$$

and choosing, as in Z. Chen [1], $\xi = X_T$, where, for a given (t_0,y_0,z_0), X is the solution of the SDE

$$X_t = y_0 - \int_{t_0}^t g_2(u, X_u, z_0)du + \int_{t_0}^t z_0\cdot dW_u,$$

we obtain, letting $t \longrightarrow t_0^+$, $g_1(t_0,y_0,z_0) \le g_2(t_0,y_0,z_0)$.

4.2 The y–independent case

Now, we turn to the case when the generator does not depend on the variable y; in this context we can answer the question asked in the introduction.
In this subsection, let $g : \Omega\times[0,T]\times\mathbb{R}^d \longrightarrow \mathbb{R}$ be a function which satisfies the assumption (A 1) and (A 3). We recall the following lemma from S. Peng [8]:

Lemma 4.2 *For each pair of random variables (ξ,η) in $\mathrm{L}^2(\mathcal{F}_T)\times\mathrm{L}^2(\mathcal{F}_t)$, we have, for $t \le r \le T$:*

$$\mathcal{E}_g\big(\xi - \eta \,|\, \mathcal{F}_r\big) = \mathcal{E}_g\big(\xi \,|\, \mathcal{F}_r\big) - \eta,$$

and, in particular, for each $0 \le s \le t \le T$, $\mathcal{E}_g\big\{\xi - \mathcal{E}_g\big(\xi \,|\, \mathcal{F}_t\big) \,|\, \mathcal{F}_s\big\} = 0$.

Actually, we can give a characterization of generators independent of y.

Lemma 4.3 *Let us assume that the assumptions (A 1), (A 3) and (A 4) hold for g. Let us suppose moreover that*

$$\forall \xi \in \mathrm{L}^2(\mathcal{F}_T),\ \forall x \in \mathbb{R},\ \forall t \in [0,T], \quad \mathcal{E}_g\big(\xi - x \,|\, \mathcal{F}_t\big) = \mathcal{E}_g\big(\xi \,|\, \mathcal{F}_t\big) - x.$$

Then g does not depend on y.

Proof. The proof is a direct application of Proposition 2.3. Indeed, if we pick a triple (t, y, z) in $[0, T[\times\mathbb{R}\times\mathbb{R}^d$, we have,

$$g(t,y,z) = \lim_{\varepsilon\to 0^+} \frac{1}{\varepsilon}\big\{\mathcal{E}_g\big(y + z\cdot(W_{t+\varepsilon} - W_t)\,\big|\,\mathcal{F}_t\big) - y\big\},$$

where the limit is taken in L^2.
On the other hand, we have $\mathcal{E}_g\big(y + z\cdot(W_{t+\varepsilon} - W_t)\,\big|\,\mathcal{F}_t\big) = y + \mathcal{E}_g\big(z\cdot(W_{t+\varepsilon} - W_t)\,\big|\,\mathcal{F}_t\big)$ which yields

$$g(t,y,z) = \lim_{\varepsilon\to 0^+} \frac{1}{\varepsilon}\mathcal{E}_g\big(z\cdot(W_{t+\varepsilon} - W_t)\,\big|\,\mathcal{F}_t\big).$$

Hence g does not depend on y. □

Let us consider, for $i = 1, 2$, $g_i : \Omega \times [0, T] \times \mathbb{R}^d \longrightarrow \mathbb{R}$ s.t. $\big(g_i(s, z)\big)_{s\in[0,T]}$ is progressively measurable for each z.

Theorem 4.4 *Let us assume that (A 1), (A 3) and (A 4) hold for g_i, $i = 1, 2$, and let us suppose that:*

$$\forall \xi \in \mathrm{L}^2(\mathcal{F}_T), \qquad \mathcal{E}_{g_1}(\xi) \le \mathcal{E}_{g_2}(\xi),$$

then, $\mathbb{P}-a.s.$, we have

$$\forall s \in [0,T],\ \forall z \in \mathbb{R}^d, \qquad g_1(s,z) \le g_2(s,z).$$

Proof. Firstly, we show that under the assumptions of the theorem, for each $\xi \in \mathrm{L}^2(\mathcal{F}_T)$ and for each $t \in [0, T]$, $\mathcal{E}_{g_1}\big(\xi \,|\, \mathcal{F}_t\big) \le \mathcal{E}_{g_2}\big(\xi \,|\, \mathcal{F}_t\big)$ and then we will apply the converse comparison theorem, Theorem 4.1 above, to get the result.
For each $\zeta \in \mathrm{L}^2(\mathcal{F}_T)$ and for each t, we have, by assumption,

$$\mathcal{E}_{g_1}\big\{\zeta - \mathcal{E}_{g_1}\big(\zeta \,|\, \mathcal{F}_t\big)\big\} \le \mathcal{E}_{g_2}\big\{\zeta - \mathcal{E}_{g_1}\big(\zeta \,|\, \mathcal{F}_t\big)\big\},$$

and by Lemma 4.2, we get $\mathcal{E}_{g_1}\big\{\zeta - \mathcal{E}_{g_1}\big(\zeta \,|\, \mathcal{F}_t\big)\big\} = 0$ and

$$\mathcal{E}_{g_2}\big\{\zeta - \mathcal{E}_{g_1}\big(\zeta \,|\, \mathcal{F}_t\big)\big\} = \mathcal{E}_{g_2}\Big(\mathcal{E}_{g_2}\big\{\zeta - \mathcal{E}_{g_1}\big(\zeta \,|\, \mathcal{F}_t\big) \,|\, \mathcal{F}_t\big\}\Big) = \mathcal{E}_{g_2}\big\{\mathcal{E}_{g_2}\big(\zeta \,|\, \mathcal{F}_t\big) - \mathcal{E}_{g_1}\big(\zeta \,|\, \mathcal{F}_t\big)\big\},$$

and thus, we deduce that

$$\forall \zeta \in \mathrm{L}^2(\mathcal{F}_T),\ \forall t \in [0,T], \qquad \mathcal{E}_{g_2}\big\{\mathcal{E}_{g_2}\big(\zeta \,|\, \mathcal{F}_t\big) - \mathcal{E}_{g_1}\big(\zeta \,|\, \mathcal{F}_t\big)\big\} \ge 0.$$

Let us fix $\xi \in \mathrm{L}^2(\mathcal{F}_T)$ and $t \in [0, T]$. Applying the previous inequality to $\zeta = \xi\mathbf{1}_{\{\eta<0\}}$, where $\eta = \mathcal{E}_{g_2}\big(\xi \,|\, \mathcal{F}_t\big) - \mathcal{E}_{g_1}\big(\xi \,|\, \mathcal{F}_t\big)$, we get, noting that $\mathcal{E}_{g_i}\big(\xi\mathbf{1}_{\{\eta<0\}} \,|\, \mathcal{F}_t\big) = \mathbf{1}_{\{\eta<0\}}\mathcal{E}_{g_i}\big(\xi \,|\, \mathcal{F}_t\big)$ since η is $\mathcal{F}_t$–measurable ([8, Lemma 36.6]), $\mathcal{E}_{g_2}\big(\eta\mathbf{1}_{\{\eta<0\}}\big) \ge 0$.

On the other hand, since obviously the random variable $\eta\mathbf{1}_{\{\eta<0\}}$ is nonpositive, we obtain from the classical comparison theorem the fact that $\mathcal{E}_{g_2}\big(\eta\mathbf{1}_{\{\eta<0\}}\big) \le \mathcal{E}_{g_2}(0) = 0$. It follows that $\mathcal{E}_{g_2}\big(\eta\mathbf{1}_{\{\eta<0\}}\big) = 0$, and by the *strict* comparison theorem (see e.g. [8, Theorem 35.3]), we deduce that $\mathbb{P}-a.s.$ $\eta\mathbf{1}_{\{\eta<0\}} = 0$ which says, coming back to the definition of the random variable η, that $\mathcal{E}_{g_2}\big(\xi\,|\,\mathcal{F}_t\big) \ge \mathcal{E}_{g_1}\big(\xi\,|\,\mathcal{F}_t\big)$, $\mathbb{P}-a.s.$
Thus, we can apply, in view of the assumption (A 3), Theorem 4.1, to obtain that, $\mathbb{P}-a.s.$

$$\forall s \in [0,T],\ \forall z \in \mathbb{R}^d, \qquad g_1(s,z) \le g_2(s,z),$$

which is the result we want. □

We will precise this result in a particular case but before we recall an elementary property of convex functions.

Lemma 4.5 *Let $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ be a convex function and let us assume that h is bounded from above. Then h is constant.*

Proof. Since h is convex, we have, for each y and for each $\alpha \in [0,1]$, $h(\alpha y) \le \alpha h(y) + (1-\alpha)h(0)$, and then, for each x and for each $\alpha \in [0,1]$, choosing $y = x/\alpha$ in the previous inequality, we get $h(x) \le \alpha h(x/\alpha) + (1-\alpha)h(0)$. Since h is upper bounded, say by M, the previous inequality yields $h(x) \le \alpha M + (1-\alpha)h(0)$, and sending α to 0^+, we get $\forall x \in \mathbb{R}^n,\quad h(x) \le h(0)$.
But, on the other hand, for each x, $2h(0) \le h(x) + h(-x)$ and by the previous inequality we deduce that $2h(0) \le h(x) + h(0)$ which says that $h(0) \le h(x)$.
Thus h is constant. □

Remark. As a byproduct of this lemma, we deduce that if g satisfies the assumptions (A 1) and (A 3) and if $(y,z) \longmapsto g(s,y,z)$ is convex then in fact g does not depend on y. Indeed, if z is fixed, the function $y \longmapsto g(s,y,z)$ is convex. Moreover, since $g(s,y,0) = 0$, the Lipschitz assumption gives $|g(s,y,z)| \le K\|z\|$ and thus by the previous lemma $y \longmapsto g(s,y,z)$ is constant. Of course, the same result holds for concave functions.

This remark explains why we only consider functions of the variable z in the following result. Keeping the same setting as in the previous theorem, we claim the corollary:

Corollary 4.6 *Under the assumptions of Theorem 4.4, if we assume moreover that, $\mathbb{P}-a.s.$, for each s, $z \longmapsto g_1(s,z)$ is convex (respectively $z \longmapsto g_2(s,z)$ is concave), then, if*

$$\forall \xi \in \mathrm{L}^2(\mathcal{F}_T), \qquad \mathcal{E}_{g_1}(\xi) \le \mathcal{E}_{g_2}(\xi),$$

there exists a progressively measurable bounded process $(\alpha_t)_{t\in[0,T]}$ s.t. $\mathbb{P}-a.s.$,

$$\forall s \in [0,T],\ \forall z \in \mathbb{R}^d, \qquad g_1(s,z) = g_2(s,z) = \alpha_s \cdot z.$$

Proof. By Theorem 4.4 above, we deduce that

$$\mathbb{P}-a.s. \quad \forall (s,z) \qquad g_1(s,z) \le g_2(s,z).$$

But the function $z \longmapsto (g_1 - g_2)(s,z)$ is convex and bounded from above by 0. Hence, we deduce, from Lemma 4.5, since $g_1(s,0) - g_2(s,0) = 0$, that $g_1(s,z) = g_2(s,z)$. It follows that $g_1(s,\cdot)$ is convex and concave and then g_1 is a linear map since $g_1(s,0) = 0$. So there exists a progressively measurable process $(\alpha_t)_{t\in[0,T]}$ s.t. $\mathbb{P}-a.s.$,

$$\forall s \in [0,T],\ \forall z \in \mathbb{R}^d, \qquad g_1(s,z) = g_2(s,z) = \alpha_s \cdot z.$$

Moreover, the process α has to be bounded since the function $g_1(s,z)$ is Lipschitz in z uniformly w.r.t s by the assumption (A 1). □

Remark. This corollary covers the case of linear generators. However, it is worth noting that linear generators are a particular case of the following situation: assume that g_i is a function which is odd in the variable z and even in the variable y. Then, it is very easy to check that $\xi \longmapsto \mathcal{E}_{g_i}(\xi)$ is also odd. Therefore, if we have, for each ξ, $\mathcal{E}_{g_1}(\xi) \le \mathcal{E}_{g_2}(\xi)$, the inequality is actually an equality. Thus by Z. Chen's result, we have $g_1 \equiv g_2$.

5 On Jensen's inequality

This short section is devoted to Jensen's inequality for g–expectation, but before we recall an elementary property of convex functions.

Lemma 5.1 *Let $h : \mathbb{R} \longrightarrow \mathbb{R}$ be a convex function. Then, for each $x \in \mathbb{R}$, we have:*

$$\begin{aligned} h(\alpha x) &\le \alpha h(x) + (1-\alpha)h(0) \quad \text{if } \alpha \in [0,1], \\ h(\alpha x) &\ge \alpha h(x) + (1-\alpha)h(0) \quad \text{if } \alpha \in]0,1[^{c}. \end{aligned}$$

Proof. Let us fix $x \in \mathbb{R}^n$ and let us set, for $\alpha \in \mathbb{R}$, $f(\alpha) = h(\alpha x) - \big(\alpha h(x) + (1-\alpha)h(0)\big)$. Then, f is a convex function on $\mathbb{R}$. Since $f(0) = f(1) = 0$, we immediately get the result. □

Now, we give a counterexample to the Jensen inequality. Consider $g : \mathbb{R} \longrightarrow \mathbb{R}$ defined as follows: $g(z) = z^2$ if $z \in [-1,1]$ and $g(z) = 2|z| - 1$ if $|z| > 1$. Let us introduce, for a fixed $\sigma \in]0,1]$, $\xi := -\sigma^2 T + \sigma W_T$. Obviously, the solution of the BSDE

$$Y_t = \xi + \int_t^T g(Z_s)ds - \int_t^T Z_s dW_s, \qquad 0 \le t \le T,$$

is $\big((-\sigma^2 t + \sigma W_t, \sigma)\big)_{t\in[0,T]}$. Hence, for each $t \in [0,T]$, we have:

$$\frac{1}{2}\mathcal{E}_g\big(\xi \,|\, \mathcal{F}_t\big) = -\frac{\sigma^2}{2}t + \frac{\sigma}{2}W_t.$$

Let us remark that, the solution of the BSDE

$$Y_t = \frac{1}{2}\xi + \int_t^T g(Z_s)ds - \int_t^T Z_s dW_s, \qquad 0 \le t \le T,$$

is given by $\Big\{\big(\frac{\sigma}{2}W_t - \frac{\sigma^2}{2}T + \big(\frac{\sigma}{2}\big)^2(T-t), \frac{\sigma}{2}\big),\ t \in [0,T]\Big\}$. It follows that:

$$\frac{1}{2}\mathcal{E}_g\big(\xi \,|\, \mathcal{F}_t\big) - \mathcal{E}_g\big(\frac{1}{2}\xi \,|\, \mathcal{F}_t\big) = (T-t)\frac{\sigma^2}{4},$$

which is positive as soon as $t < T$.

This contradicts Jensen's inequality in the simplest case i.e. the function is linear ($x \longmapsto x/2$) and the generator is convex which implies, by the comparison theorem, that $\xi \longmapsto \mathcal{E}_g(\xi)$ is convex.

We end this paragraph by studying the Jensen inequality in a particular case: g independent of y and g convex in z.

We assume moreover that $\mathbb{P}-a.s.$, $z \longmapsto g(s,z)$ is convex for each s, and we fix $\xi \in \mathrm{L}^2(\mathcal{F}_T)$. Let us consider a function $F:\mathbb{R} \longrightarrow \mathbb{R}$ which is convex and s.t. the random variable $F(\xi)$ is square integrable. By $\partial F(x)$ we mean the subdifferential of the convex function F at the point x (see e.g. R. T. Rockafellar [11]) which, in our context, reduces to the interval $[F'_g(x), F'_d(x)]$.

Proposition 5.2 *If $\mathbb{P}-a.s.$, $\partial F\big(\mathcal{E}_g(\xi \,|\, \mathcal{F}_t)\big) \cap]0,1[^{\mathrm{c}} \neq \emptyset$, then*

$$\mathbb{P}-a.s. \qquad F\big(\mathcal{E}_g(\xi \,|\, \mathcal{F}_t)\big) \leq \mathcal{E}_g\big(F(\xi) \,|\, \mathcal{F}_t\big).$$

Proof. By the assumption on the subdifferential of F, we have

$$F(\xi) - F\big(\mathcal{E}_g(\xi \,|\, \mathcal{F}_t)\big) \geq \beta\big(\xi - \mathcal{E}_g(\xi \,|\mathcal{F}_t)\big),$$

where β is an $\mathcal{F}_t$–measurable random variable with values in $]0,1[^{\mathrm{c}}$. Let us a pick a positive integer n. Multiplying the previous inequality by $\mathbf{1}_{\Omega_n}$ where Ω_n stands for the set $\big\{|\mathcal{E}_g(\xi|\,\mathcal{F}_t)| + |\beta| \leq n\big\}$ we have, setting moreover $\beta_n = \beta \mathbf{1}_{\Omega_n}$,

$$\mathbf{1}_{\Omega_n}\big[F(\xi) - F\big(\mathcal{E}_g(\xi \,|\, \mathcal{F}_t)\big)\big] \geq \beta_n\big(\xi - \mathcal{E}_g(\xi \,|\mathcal{F}_t)\big).$$

It is worth noting that the four terms of the previous inequality are square integrable since ξ, $F(\xi)$ belong to $\mathrm{L}^2(\mathcal{F}_T)$, β_n is bounded by n and F is continuous on $\mathbb{R}$ as a convex function. Thus, taking into account the fact that $\mathcal{E}_g(\cdot \,|\, \mathcal{F}_t)$ is nondecreasing, we deduce that

$$\mathcal{E}_g\Big(\mathbf{1}_{\Omega_n}\big[F(\xi) - F\big(\mathcal{E}_g(\xi \,|\, \mathcal{F}_t)\big)\big] \,\big|\, \mathcal{F}_t\Big) \geq \mathcal{E}_g\Big(\beta_n\big(\xi - \mathcal{E}_g(\xi \,|\mathcal{F}_t)\big) \big|\, \mathcal{F}_t\Big)$$

and since β_n and $\mathbf{1}_{\Omega_n}$ are $\mathcal{F}_t$–measurable, Lemma 4.2 implies

$$\mathcal{E}_g\big(\mathbf{1}_{\Omega_n}F(\xi) \,|\, \mathcal{F}_t\big) - \mathbf{1}_{\Omega_n}F\big(\mathcal{E}_g(\xi \,|\, \mathcal{F}_t)\big) \geq \mathcal{E}_g\big(\beta_n\xi \,|\, \mathcal{F}_t\big) - \beta_n\mathcal{E}_g\big(\xi \,|\, \mathcal{F}_t\big). \tag{9}$$

On the other hand, since β_n is $\mathcal{F}_t$–measurable and since $\beta_n \in]0,1[^{\mathrm{c}}$, we have the inequality $\mathcal{E}_g\big(\beta_n\xi \,|\, \mathcal{F}_t\big) \geq \beta_n\mathcal{E}_g\big(\xi \,|\, \mathcal{F}_t\big)$. Indeed, for $t \leq r \leq T$,

$$\beta_n Y_r = \beta_n\xi + \int_r^T \beta_n g(s, Z_s)ds - \int_r^T \beta_n Z_s \cdot dW_s,$$

and since g is convex and $\beta_n \in]0,1[^{\mathrm{c}}$, Lemma 5.1 shows that, using the fact that $g(s,0) = 0$, for each $t \leq r \leq T$,

$$\beta_n Y_r \leq \beta_n\xi + \int_r^T g(s, \beta_n Z_s)ds - \int_r^T \beta_n Z_s \cdot dW_s.$$

As a byproduct of this inequality, we get that $\beta_n\mathcal{E}_g\big(\xi \,|\, \mathcal{F}_t\big) \leq \mathcal{E}_g\big(\beta_n\xi \,|\, \mathcal{F}_t\big)$, and coming back to the inequality (9), we obtain the inequality

$$\forall n \in \mathbb{N}^*, \quad \mathbf{1}_{\Omega_n}F\big(\mathcal{E}_g(\xi \,|\, \mathcal{F}_t)\big) \leq \mathcal{E}_g\big(\mathbf{1}_{\Omega_n}F(\xi) \,|\, \mathcal{F}_t\big).$$

If $n \to +\infty$, $\mathbf{1}_{\Omega_n}F(\xi)$ converges to $F(\xi)$ in $\mathrm{L}^2(\mathcal{F}_T)$ and since $\xi \longmapsto \mathcal{E}_g(\xi \,|\, \mathcal{F}_t)$ is a continuous map from $\mathrm{L}^2(\mathcal{F}_T)$ into $\mathrm{L}^2(\mathcal{F}_t)$ (see e.g. S. Peng [8, Lemma 36.9])we conclude the proof of the Jensen inequality in this context letting n tends to infinity. □

Remark. It is also possible to obtain the Jensen inequality when the function g is concave instead of being convex. In this context, this requires essentially the fact that the convex function F satisfies $0 \leq F'(x) \leq 1$ (e.g. $F(x) = 1/\pi\{x\mathrm{Arctan}x - 1/2\ln(1+x^2)\} + x/2$).

References

[1] Z. Chen, *A property of backward stochastic differential equations*, C. R. Acad. Sci. Paris Sér. I Math. **326** (1998), no. 4, 483–488.

[2] N. El Karoui, S. Peng, and M.-C. Quenez, *Backward stochastic differential equations in finance*, Math. Finance **7** (1997), no. 1, 1–71.

[3] H. Kunita, *Stochastic differential equations and stochastic flows of diffeomorphisms*, Ecole d'été de probabilités de Saint-Flour XII—1982 (P.-L. Hennequin, ed.), Lecture Notes in Math., vol. 1097, Springer-Verlag, Berlin Heidelberg New York, 1984, pp. 143–303.

[4] E. Pardoux and S. Peng, *Adapted solution of a backward stochastic differential equation*, Systems Control Lett. **14** (1990), no. 1, 55–61.

[5] E. Pardoux and S. Peng, *Backward stochastic differential equations and quasilinear parabolic partial differential equations*, Stochastic partial differential equations and their applications (Charlotte, NC, 1991) (B. L. Rozovskii and R. B. Sowers, eds.), Lecture Notes in Control and Inform. Sci., vol. 176, Springer, Berlin, 1992, pp. 200–217.

[6] S. Peng, *Probabilistic interpretation for systems of quasilinear parabolic partial differential equations*, Stochastics Stochastics Rep. **37** (1991), no. 1-2, 61–74.

[7] S. Peng, *Stochastic Hamilton-Jacobi-Bellman equations*, SIAM J. Control Optim. **30** (1992), no. 2, 284–304.

[8] S. Peng, *Backward SDE and related g–expectation*, Backward stochastic differential equations (N. El Karoui and L. Mazliak, eds.), Pitman Res. Notes Math. Ser., vol. 364, Longman, Harlow, 1997, pp. 141–159.

[9] S. Peng, *Monotonic limit theorem of BSDE and nonlinear decomposition theorem of Doob–Meyer's type*, Probab. Theory Related Fields **113** (1999), no. 4, 23–30.

[10] F. Pradeilles, *Wavefront propagation for reaction-diffusion systems and backward SDEs*, Ann. Probab. **26** (1998), no. 4, 1575–1613.

[11] R. T. Rockafellar, *Convex analysis*, Princeton Math. Ser., vol. 28, Princeton University Press, Princeton, N.J., 1970.